BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 6014–6018, August 1990
Developmental Biology

Human homolog of the mouse sperm receptor

(human zona pellucida/*ZP3*/oocyte-specific gene expression/cis-acting elements/human sperm receptor)

MARGARET E. CHAMBERLIN AND JURRIEN DEAN

Laboratory of Cellular and Developmental Biology, National Institute of Arthritis, Diabetes, and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892

Communicated by Christian B. Anfinsen, May 31, 1990

ABSTRACT **The human zona pellucida, composed of three glycoproteins (ZP1, ZP2, and ZP3), forms an extracellular matrix that surrounds ovulated eggs and mediates species-specific fertilization. The genes that code for at least two of the zona proteins (ZP2 and ZP3) cross-hybridize with other mammalian DNA. The recently characterized mouse sperm receptor gene (*Zp-3*) was used to isolate its human homolog. The human homolog spans ≈18.3 kilobase pairs (kbp) (compared to 8.6 kbp for the mouse gene) and contains eight exons, the sizes of which are strictly conserved between the two species. Four short (8–15 bp) sequences within the first 250 bp of the 5′ flanking region in the human *Zp-3* homolog are also present upstream of mouse *Zp-3*. These elements may modulate oocyte-specific gene expression. By using the polymerase chain reaction, a full-length cDNA of human ZP3 was isolated from human ovarian poly(A)$^+$ RNA and used to deduce the structure of human ZP3 mRNA. Certain features of the human and mouse ZP3 transcripts are conserved. Both have unusually short 5′ and 3′ untranslated regions, both contain a single open reading frame that is 74% identical, and both code for 424 amino acid polypeptides that are 67% the same. The similarity between the two proteins may define domains that are important in maintaining the structural integrity of the zona pellucida, while the differences may play a role in mediating the species-specific events of mammalian fertilization.**

Many details of mammalian fertilization have been elucidated from studies in the mouse. Fertilization begins when a capacitated mouse spermatozoon attaches to the zona pellucida, an extracellular matrix that surrounds the ovulated egg of mammals. After tight binding to the zona, the dehiscence of the sperm acrosome results in the release of lytic enzymes that, coupled with the forward motility of the sperm, results in passage through the zona pellucida. The sperm then traverses the perivitelline space, fuses with the egg's plasma membrane, and is absorbed into the cytoplasm, where its nucleus undergoes decondensation and repackaging into the male pronucleus of the one-cell zygote (1).

The sperm–egg interaction that takes place in the oviduct after ovulation of the egg and insemination of sperm is relatively species-specific (2–5). In part, this specificity is dependent on the affinity of the spermatozoon for the zona pellucida of the homologous species (for review, see ref. 6). The mouse zona pellucida is composed of three sulfated glycoproteins, ZP1, ZP2, and ZP3, and specific functions have been ascribed to each. ZP3 induces the sperm acrosome reaction and mediates the initial binding of sperm to the egg via O-linked oligosaccharide side chains. ZP2 acts as a secondary sperm receptor and it, along with ZP3, is biochemically modified after fertilization to provide the postfertilization block to polyspermy. ZP2 and ZP3 exist as dimers in long filaments that appear to be cross-linked by ZP1 (for review, see ref. 7). The recent cloning of mouse ZP2 and ZP3 cDNAs (8–10) and the characterization of their genomic loci (8, 11, 12) has provided a wealth of molecular detail on the primary structure of the zona proteins and their developmentally regulated expression during oogenesis (8, 13).

Less is known about the primary structure and the functions of the zona pellucida proteins from other species, including that of human. The human zona pellucida is composed of three glycoproteins, ZP1 (90–110 kDa), ZP2 (64–76 kDa), and ZP3 (57–73 kDa) (14, 15), at least one of which appears to be modified following fertilization (15). However, more detailed biochemical studies of individual human zona pellucida proteins are difficult because of the paucity of biological material. An alternate approach to learning more about the human zona pellucida is to deduce the structure of individual proteins from their cognate genes. We have demonstrated (9, 10) that the genes that code for the zona pellucida proteins are conserved among mammals. Taking advantage of the cross-hybridization of a mouse cDNA with human DNA, we now report the isolation of full-length cDNA clones of human ZP3* and the characterization of the genomic locus of human *ZP3*, the homologue to the mouse sperm receptor.

MATERIALS AND METHODS

Human Genomic DNA. A human genomic library in Charon 4A (16) was screened with pZP3.2 (10). A recombinant, λHuZP3.14, was isolated and plaque-purified. The 14.0-kilobase-pair (kbp) insert (containing exons 1–5) was digested with *Bam*HI and *Eco*RI, the resultant fragments were subcloned into pBluescript, and the nucleic acid sequence of 11.7 kbp was determined (17).

The 3′ end of the gene including exons 6–8 was obtained by the polymerase chain reaction (PCR) (GeneAmp; Perkin–Elmer/Cetus). One microgram of human genomic DNA from K562 cells (gift of A. Dean) was primed with oligonucleotide H78 (exon 6) and H77 (exon 8) for 25 cycles (94°C for 1 min, 60°C for 2 min, and 72° for 3 min) and a final extension at 72°C for 15 min. The 1.8-kbp fragment was blunt end–ligated into pBluescript (Stratagene) and sequenced (17).

The 3′ flanking sequences were determined by PCR of a circularized *Xba* I fragment encompassing exon 8 (Fig. 1). One microgram of human genomic DNA was digested with *Xba* I, extracted, precipitated with ETOH, resuspended (1 μg/ml), and cyclized overnight with 10 units of T4 ligase. The *Xba* I fragment containing exon 8 and the 3′ flanking region were amplified by using exon 8 oligonucleotide primers H101 and H86 for 25 cycles of PCR. The resultant 570-bp fragment was blunt end-ligated into pBluescript and sequenced (17).

Southern Blot Analysis. Exon-specific DNA probes were amplified by PCR with human cDNA as a substrate (see below) and the following oligonucleotide primer sets: H36

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: PCR, polymerase chain reaction; nt, nucleotides.
*The sequence reported in this paper has been deposited in the GenBank data base (accession no. M35109).

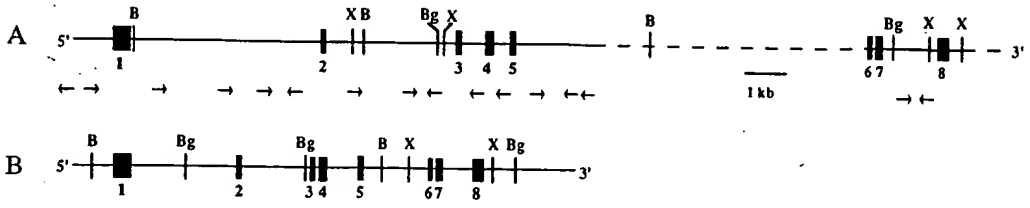


FIG. 1. Genomic locus of the human homolog of mouse *Zp-3*. (*A*) Exon map of human *Zp-3* homolog with eight exons (1–8) and restriction enzyme sites for *Bam*HI (B), *Bgl* II (Bg), and *Xba* I (X). Horizontal arrows indicate *Alu* repeat sequences. Solid lines indicate regions for which the nucleic acid sequence was determined. Distance of the dashed line was estimated from Southern blot analysis. (*B*) Exon map of mouse *Zp-3* (12) for comparison.

and H75 (exons 4 and 5); H78 and H81 (exons 6 and 7); and H80 and H77 (exon 8). Labeled probes (18) were used in Southern blots of human K562 DNA after digestion with a variety of restriction endonucleases as described (12).

Human ZP3 cDNA. Poly(A)$^+$ RNA (19) was purified from total RNA (20) isolated from a human ovary (supplied by the National Disease Research Interchange) and used as a template for first-strand synthesis with oligonucleotide primer A2T15 according to the manufacturer's instruction (Promega Riboclone cDNA synthesis system). The first strand was amplified by PCR with A2T15 and oligonucleotide H36 from human exon 4 (94°C for 5 min, 50°C for 2 min, and 72°C for 40 min, followed by 40 cycles of 94°C for 2 min, 58°C for 2 min, and 72°C for 3 min) (21). The resultant 800-bp band was reamplified with A2 [lacking poly (dT)] and H74 from exon 5 (25 cycles of 94°C for 1 min, 55°C for 2 min, and 72°C for 3 min and a final extension at 72°C for 15 min). The 670-bp insert was blunt-end ligated into pBluescript and sequenced (17). The full-length cDNA was prepared by using oligonucleotide primers H87 (exon 1) and H88 (exon 8) and 25 cycles of PCR (94°C for 1 min, 60°C for 2 min, and 72°C for 3 min and a final extension at 72°C for 15 min).

Oligonucleotide Primers. Oligonucleotide primers were synthesized on an Applied Biosystems DNA Synthesizer (model 380B). Oligonucleotide primers (5′ to 3′) hybridizing to the coding strand include: A2 (CTCGAGAAGCTTGTCGAC), A2T15 (CTCGAGAAGCTTGTCGACT$_{15}$), H77 (AAGCAGACACAGGGTGG), H81 (GCCTGCGGTTACGGGAA), H87 (AGATCTGAGCTCATTGCTTTCTTCTTTTATTCGGAAG), and H101 (GGAAGATCAGTGGCCCC). Oligonucleotide primers (5′ to 3′) hybridizing to the noncoding strand include: H36 (TGCAGCCCACCTCCAGG), H74 (CTGTCTTGTCGACGGTC), H78 (TCACCTGCCACCTGAAG), H80 (CAGAAGAAGCAGATGTC), H86 (CTGTGGTGGTGTCCCTG), and H88 (TGCAGGGTACCATGGAGCTGAGCTATAGGC).

Computer-Aided Sequence Analysis. Sequence data collection and comparison of nucleic acid sequences was performed with Microgenie Software (Beckman, version 6). Comparison of the human and mouse polypeptides allowed for conservative amino acid substitutions (22). Secondary structure of the deduced amino acid sequence of human ZP3 was determined (23) as was its hydropathicity (24) and hydrophilicity (25). Signal peptide cleavage sites were defined by a sliding window-weighted matrix algorithm (26).

RESULTS

Genomic Locus of Human *Zp-3* Homolog. A clone containing a 14.0-kbp insert was isolated from a human genomic library (16) screened with a mouse ZP3 cDNA probe (10). Of the 14.0-kbp insert, 11.7 kbp were sequenced and contained the first five exons of the human *Zp-3* homolog as well as ≈1.1 kbp of the 5′ flanking region and 1.78 kbp of intron 5 (27). Despite repeated screening of the library with probes from mouse cDNA specific for exons 6–8, a genomic fragment containing the balance of the human exons was not isolated.

Instead, the remaining exons of the human *Zp-3* homolog were characterized by amplifying genomic DNA by the PCR with oligonucleotides from exons 6 and 8 (based on the cDNA sequence). The resultant 1.8-kbp fragment containing exons 6–8 and intron 7 was sequenced. The size of the remaining intron, 5, was determined as follows. Radiolabeled probes corresponding to exons 4, 5, 6, and 7 all hybridized to a single 10.5-kbp *Xba* I fragment from human genomic DNA (an exon 8 probe detected a 0.6-kbp fragment). Sequence analysis showed that exon 5 was 1521 bp 3′ of a *Xba* I site and that exon 6 was 1460 bp 5′ of a *Xba* I site in intron 7. Thus, the limits of intron 5 can be set at 7.5 kbp ± 10%.

The resultant exon map of the human *Zp-3* homolog (Fig. 1*A*) is similar to that of mouse *Zp-3* (Fig. 1*B*) (11, 12). The size of the human exons 1–8 range from 92 to >312 bp (Table 1): exons 2, 3, 5, 6, and 7 are identical in length to those of the mouse gene, whereas exon 4 is 6 bp smaller and exon 8 is 4 bp larger than the corresponding mouse exons. The sizes of the introns in the human gene (130 bp to ≈7.5 kb) vary considerably from those in mouse *Zp-3*, although the clustering of exons 3–5 and 6–8 appears to be similar in both (Fig. 1 and Table 1). Overall, the 18.3-kbp transcription unit of the human *Zp-3* homolog is more than twice the 8.6-kbp size of the mouse gene (11, 12).

The exact determination of the transcription start site was precluded by the paucity of human ovarian RNA. However, there is a TATAA box 43 bp upstream of the ATG (compared

Table 1. Exon sizes, intron lengths, and splice-junction sequences of human *Zp-3* homolog

Number	Exon size, bp	Splice-junction sequences	Intron length, bp
1	>330	.....Exon 1....*GT*AAGAGAGGCT	4260
2	119	CCTTTCCTCC*AG*....Exon 2....*GT*CGGTGTGGGA	2813
3	104	TTCTTCTCTC*AG*....Exon 3....*GT*AAGAGAAGAA	449
4	178	TGTGTCTTTC*AG*....Exon 4....*GT*GAGCACTGGG	389
5	118	TTTTCTTCCA*AG*....Exon 5....*GT*AAGAGCTTTA	7500 ± 10%
6	92	Not defined ....Exon 6....*GT*GAGGACAGGT	130
7	137	CTTTCATCTC*AG*....Exon 7....*GT*ATGTCACAGA	1234
8	229	CTCCTTTCAC*AG*....Exon 8.....	—

The italicized letters indicate splice acceptor and donor consensus sequences.

to 58 bp in the mouse gene) that makes it likely that the 5′ untranslated region is similar in length to the 29 nucleotides (nt) of the 5′ untranslated region of the mouse ZP3 mRNA. The 3′ end of exon 8 was determined from the nucleic acid sequence of two separate amplified isolates of human cDNAs (see below) and two separate amplified isolates of human genomic DNA.

Sequence Environment of Human *Zp-3* Homolog. Because of the inexactitude of the transcription start site of the human gene, the 5′ flanking sequences of the mouse *Zp-3* gene and the human homolog were aligned on their TATAA boxes; in the mouse, the TATAA box is 30 bp upstream of the transcription start site (Fig. 2). Human *Zp-3* homolog contains two *Alu* repeats (28) inverted one to the other (−254 to −608; −784 to −1111). The right-hand monomer of the repeat closest to the coding region of the gene corresponds to the homologous B1 repeat (29) found upstream of mouse *Zp-3* (12). Although present in the genome on average once every 5–10 kbp (28), there are at least 16 human *Alu* repeat sequences present at the locus of the human *Zp-3* homolog (Fig. 1). The tandem repeat found upstream of the transcription start site of mouse *Zp-3* and the perfect 12-bp repeat found 3′ to the gene (12) are not present in the human gene.

There is little sequence similarity in the immediate 5′ flanking regions of the mouse *Zp-3* gene and its human homolog, and neither contains a canonical CCAAT box. However, within the first 250 bp of the human homolog and the first 200 bp of 5′ flanking sequence of the mouse *Zp-3* gene there are four elements that range in length from 8 to 15 bp and are 82–90% identical (Fig. 2). Elements 1 and 2 are very similarly positioned and, although elements 3 and 4 are located more 5′ in the human than in the mouse sequence, the distances between the two are virtually identical (34 and 35 bp, respectively).

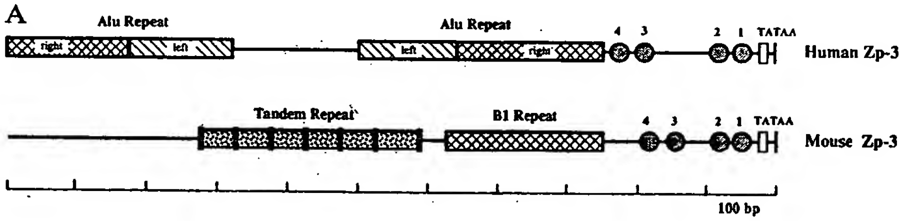
Deduced Structure of Human ZP3 mRNA and Protein. Several cDNAs were obtained as PCR products using poly(A)$^+$ RNA isolated from human ovarian tissue as a substrate and an oligo(dT) adaptor primer in conjunction with oligonucleotide primers from exons 1, 5, or 6 of the human *Zp-3* homolog. Three independent amplified products were sequenced in their entirety. The data that corresponded to the genomic sequence (see above) were used to deduce the structure of human ZP3 mRNA.

The human ZP3 mRNA, devoid of its poly(A) tail, is ≈1.3 kb long with a single open reading frame of 1272 nt (Fig. 3) and has a nucleic acid sequence 74% identical to that of mouse ZP3. The initiator ATG is located in the consensus sequence ANNATG (30, 31), and the TAA stop codon is part of the polyadenylylation signal (AATAAA) that precedes the start of the poly(A) tail by 17 nt. The unusually short 5′ and 3′ untranslated regions are similar to those of mouse ZP3 mRNA (10) and of mouse ZP2 mRNA (8).

The single open reading frame encodes a 424-amino acid protein (12% acidic, 8% basic, 7% aromatic, and 32% hydrophobic residues) that is identical in length with the mouse ZP3 protein and has a calculated molecular mass of 47,032 Da (Fig. 3). Overall, 67% of the amino acids are identical between the human and mouse ZP3 proteins, with the greatest similarity in the center of the protein (Fig. 3). The human polypeptide chain contains four potential N-linked glycosylation sites, three of which are conserved in the mouse protein, including one that has been shown to be derivatized (9, 10). There are 66 potential O-linked glycosylation sites (threonines or serines), 71% of which are conserved in the mouse. The zona proteins are secreted, and, using the sliding window/matrix scoring method of von Heijne (26), we have identified a potential peptidase cut site after the 22nd amino acid that would result in a N-terminal glutamine as we have proposed for the mouse ZP3 protein (10). The secreted protein would have a core molecular mass of 44,399 Da.

Using the primary amino acid sequence, we compared the predicted secondary structure of the human and mouse proteins (Fig. 4). The hydropathicity profiles (24) of the two proteins (Fig. 4*A*) are strikingly similar and represent the conservative nature of the allowed amino acid substitutions. The same broad hydrophobic region near the carboxyl terminus observed in mouse ZP2 and ZP3 (8, 10) is also present in human ZP3. The greatest differences in hydropathicity occur near the N terminus (Fig. 4*A*) and contain a region of the proteins where the amino acid sequence is dissimilar (<40% identity). Predicted α-helical structure (23) is also well conserved (Fig. 4*B*), again reflecting the similarity between the two amino acid sequences. Both of the proteins are known to contain disulfide bonds, and interestingly, all 13 cysteine residues found in the secreted human ZP3 polypeptide are conserved in the mouse protein, suggesting a role in

A



B

Gene	Element	Sequence	Identity
Human Zp-3 Mouse Zp-3	TATAA Box	-30 [illegible] -23 -30 [illegible] -23	100%
Human Zp-3 Mouse Zp-3	Element 1	-58 [illegible]C[illegible]AG[illegible] -44 -55 [illegible]G[illegible]GT[illegible] -41	90%
Human Zp-3 Mouse Zp-3	Element 2	-85 [illegible]G[illegible]C[illegible] -73 -73 [illegible]A[illegible]A[illegible] -61	85%
Human Zp-3 Mouse Zp-3	Element 3	-206 [illegible]C[illegible] -199 -150 [illegible]A[illegible] -143	88%
Human Zp-3 Mouse Zp-3	Element 4	-240 [illegible]TA[illegible] -230 -185 [illegible]AT[illegible] -175	82%

FIG. 2. Conservation of 5′ flanking DNA sequences of mouse *Zp-3* and the human homolog. (*A*) Schematic representation of 1100 bp upstream of the transcription start site with the sequence of the human *Zp-3* homolog aligned on the TATAA box of the previously reported mouse gene (12). Numbered circles represent conserved elements (1–4) defined in *B*. Boxes represent blocks of repeated sequences found in the *Zp-3* loci. (*B*) DNA elements in human and mouse conserved in sequence and position upstream of the *Zp-3* gene. Shaded areas indicate nucleic acid identities.

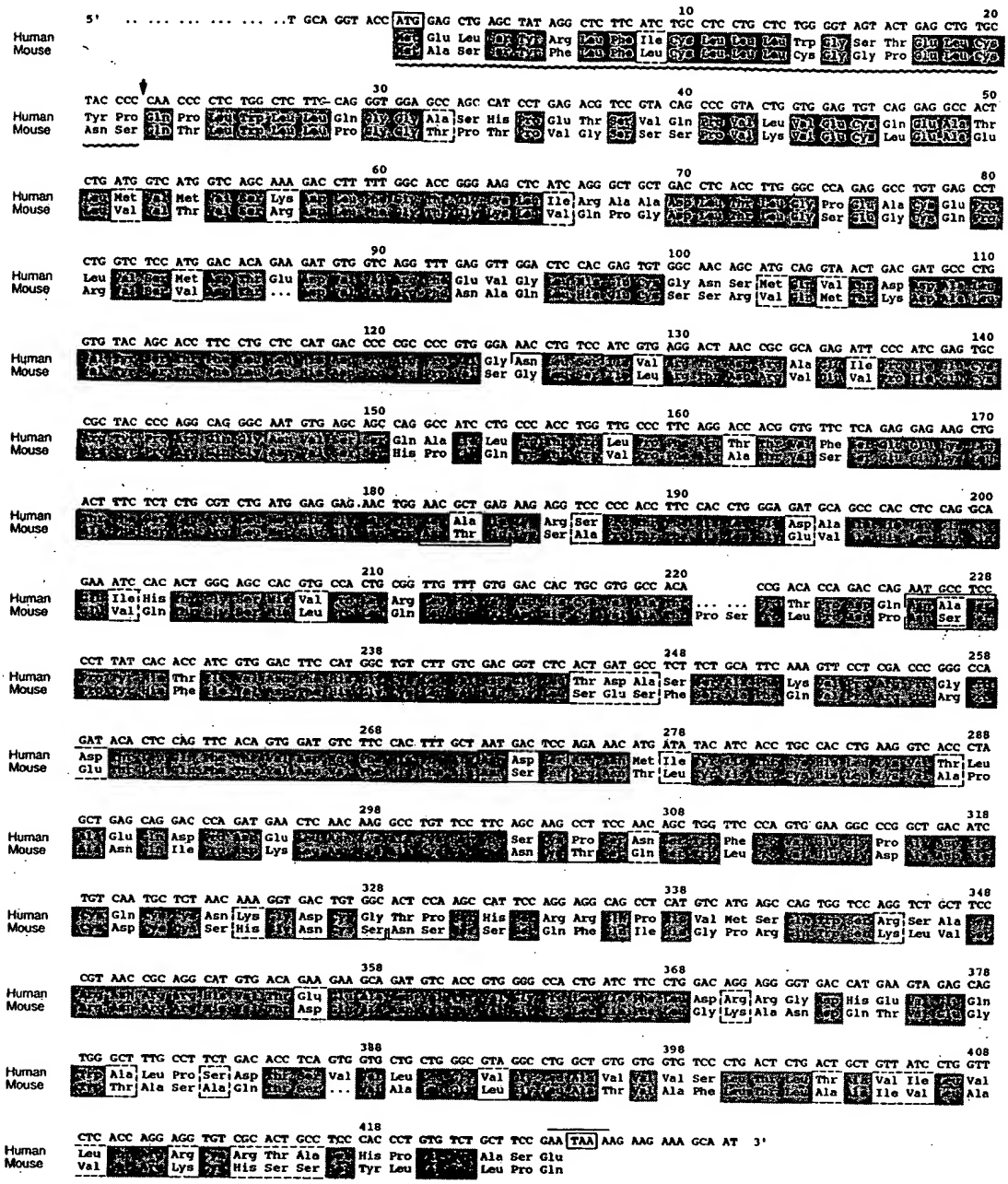


FIG. 3. Structure of human ZP3 mRNA and protein. The first line is the nucleic acid sequence of human ZP3 mRNA containing 1289 nt determined from human cDNA and genomic sequences. The initiation and stop codons are boxed, and the polyadenylylation signal is overlined. The single 1272-nt open reading frame is translated into a 424-amino acid peptide in the second line and aligned in the third line with the 424 amino acids of mouse ZP3 protein (10). Identical amino acid residues in mouse and human ZP3 are shaded; conserved changes (22) are enclosed in boxes with dotted lines. The putative 22-amino acid signal peptide is indicated by a wavy line, and the arrow points to the proposed signal peptidase cut site. The four potential N-linked glycosylation sites [Asn-Xaa-(Thr or Ser)] of human ZP3 are bracketed from above, and the six potential sites of mouse ZP3 are bracketed from below. Three of the sites are conserved between the two species.

preserving the secondary structure of these structural proteins.

DISCUSSION

The single copy human *ZP3* gene is composed of eight exons in a transcription unit of 18.3 kb. The exons, ranging in size from 92 to at least 312 bp, are almost identical in size to the eight exons of mouse *Zp-3*, and the nucleic acid sequence of the coding regions is 74% identical. We isolated a full-length cDNA from human ovarian poly(A)$^+$ even though the small amount of biological tissue available to us precluded the detection of ZP3 transcripts by RNA blot-hybridization (Northern) analysis, despite prolonged exposure times. In the mouse, *Zp-3* is transcribed uniquely in oocytes during a narrow 2-week growth phase prior to meiotic maturation and ovulation (8, 13). Presumably, the temporal and spatial specificity of this expression is modulated by transcriptional factors (32) that interact with the 5′ flanking regions of the zona genes. Binding sites for such factors may be conserved among mammals and four similar elements (8–15 bp, 82–90% identical; see Table 1) are present at approximately the same

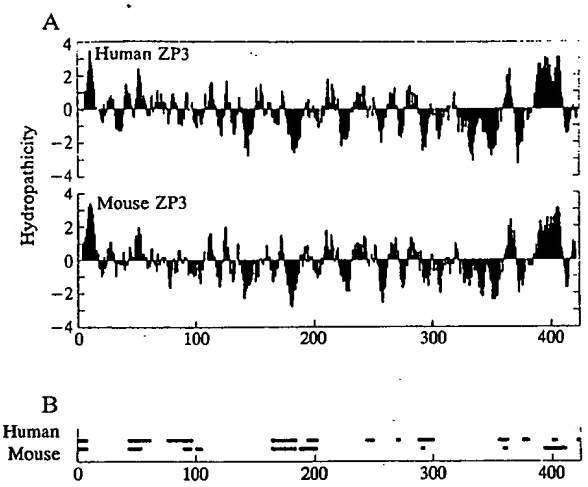


FIG. 4. Comparison of the secondary structure of human and mouse ZP3 proteins. (*A*) Hydropathicity of human and mouse ZP3 determined by the Kyte and Doolittle algorithm (24), indicating the degree of conservation between the two proteins. (*B*) α-Helical structure of human and mouse ZP3 as determined by Garnier *et al.* (23).

location in the first 250 bp upstream of the mouse *Zp-3* gene and the human homolog. These sequences are distinct from cis-acting elements previously reported in the literature. Provocatively, these four elements are also present upstream of the transcription start site of the mouse *Zp-2* gene (8, 27). Whether or not these DNA sequences have a functional role in oocyte-specific gene expression remains to be determined.

The human ZP3 transcript is remarkably similar to the mouse ZP3 mRNA. Both have short 5′ and 3′ untranslated regions and both have a single open reading frame of 1272 nt that encodes a 424-amino acid protein. Although the molecular size of the polyadenylylated human transcript is not known, the size of three other ZP3 transcripts (mouse, rat, and rabbit) are indistinguishable from one another (10), which suggests that these motifs may be common to all mammalian ZP3 mRNAs. The deduced amino acid sequences of the human and mouse ZP3 protein are 67% identical, and, as evidenced by similarity of their predicted secondary structure, many substitutions are conservative. Both the conserved central regions and the carboxyl hydrophobic domain of the human and mouse ZP3 may be important for interactions with ZP2 and ZP1, and the two other zona proteins of the mouse and human zonae pellucidae.

The similarities of the ZP3 proteins among species may be important in providing structural integrity to the extracellular zona pellucida and in positioning protein and carbohydrate moieties so that they can participate in sperm–egg interactions. If, as is the case in the mouse, human ZP3 functions as the sperm receptor *in vivo*, the regions of maximal differences may play a role mediating high-affinity interactions that lead to fertilization for a particular species. Of the four regions of the human ZP3 protein that are most dissimilar (≤40%) from the mouse protein (Fig. 3, amino acids 32–40, 93–102, 323–341, and 369–378), the two most carboxyl regions correspond to hydrophilic peaks (25) that presumably lie on the surface of the protein. One includes amino acids 338–342 that corresponds to the binding site on mouse ZP3 (33) of an anti-mouse ZP3 monoclonal antibody (34) that further supports a surface position for this domain. The cloning of two sperm receptors, each in a species for which *in vitro* fertilization is possible, may provide the necessary reagents with which to gain additional insights into those protein and carbohydrate domains that are critical for species-specific fertilization.

We appreciate the critical reading of the manuscript by Drs. Graeme Wistow and Michael O'Rand. The human ovarian tissue was kindly supplied by the National Disease Research Interchange, Philadelphia.

1. Yanagimachi, R. (1988) in *The Physiology of Reproduction*, eds. Knobil, E. & Neil, J. (Raven, New York), pp. 135–185.
2. Bedford, J. M. (1977) *Anat. Rec.* **188**, 477–488.
3. Schmell, R. D. & Gulyas, B. J. (1980) *Biol. Reprod.* **23**, 1075–1085.
4. Swenson, C. E. & Dunbar, B. S. (1982) *J. Exp. Zool.* **219**, 97–104.
5. O'Rand, M. G., Matthews, J. E., Welch, J. E. & Fisher, S. J. (1985) *J. Exp. Zool.* **235**, 423–428.
6. O'Rand, M. G. (1988) *Gamete Res.* **19**, 315–328.
7. Wassarman, P. M. (1988) *Annu. Rev. Biochem.* **57**, 415–442.
8. Liang, L.-F., Chamow, S. M. & Dean, J. (1990) *Mol. Cell. Biol.* **10**, 1507–1515.
9. Ringuette, M. J., Sobieski, D. A., Chamow, S. M. & Dean, J. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 4341–4345.
10. Ringuette, M. J., Chamberlin, M. E., Baur, A. W., Sobieski, D. A. & Dean, J. (1988) *Dev. Biol.* **127**, 287–295.
11. Kinloch, R. A., Roller, R. J., Fimiani, C. M., Wassarman, D. A. & Wassarman, P. M. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 6409–6413.
12. Chamberlin, M. E. & Dean, J. (1989) *Dev. Biol.* **131**, 207–214.
13. Philpott, C. C., Ringuette, M. J. & Dean, J. (1987) *Dev. Biol.* **121**, 568–575.
14. Sacco, A. G., Yurewicz, E. C., Subramanian, M. G. & DeMayo, F. (1981) *Biol. Reprod.* **25**, 997–1008.
15. Shabanowitz, R. B. & O'Rand, M. G. (1988) *J. Reprod. Fertil.* **82**, 151–161.
16. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. (1978) *Cell* **15**, 1157–1174.
17. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
18. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13.
19. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 1408–1412.
20. Chirgwin, J. M., Przybla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
21. Frohman, M. A., Dush, M. K. & Martin, G. R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8998–9002.
22. Dayhoff, M. O. (1969) *Altas of Protein Sequence and Structure* (Natl. Biomed. Res. Found., Washington, DC).
23. Garnier, J., Osguthorpe, D. J. & Robson, B. (1978) *J. Mol. Biol.* **120**, 97–120.
24. Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105–132.
25. Hopp, T. P. & Woods, K. P. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 3824–3828.
26. Von Heijne, G. (1986) *Nucleic Acids Res.* **14**, 4683–4690.
27. Chamberlin, M. E. (1989) Ph.D. Dissertation (The Johns Hopkins Univ., Baltimore).
28. Schmid, C. W. & Jelinek, W. R. (1982) *Science* **216**, 1065–1070.
29. Krayev, A. S., Kamerov, D. A., Skyrabin, K. G., Ryskor, A. P., Bayev, A. A. & Georgiev, G. P. (1980) *Nucleic Acids Res.* **8**, 1201–1215.
30. Kozak, M. (1986) *Cell* **44**, 283–292.
31. Cavener, D. R. (1986) *Nucleic Acids Res.* **15**, 1353–1361.
32. Mitchell, P. J. & Tjian, R. (1989) *Science* **245**, 371–378.
33. Millar, S. E., Chamow, S. M., Baur, A. W., Oliver, C., Robey, F. & Dean, J. (1989) *Science* **246**, 935–938.
34. East, I. J., Gulyas, B. J. & Dean, J. (1985) *Dev. Biol.* **109**, 268–273.

AUGUST 1990
VOLUME 87
NUMBER 16

Proceedings OF THE National Academy of Sciences

OF THE UNITED STATES OF AMERICA